Structure Preserving Algorithms for Perplectic Eigenproblems*

D. Steven Mackey[†] Niloufer Mackey[‡] Daniel M. Dunlavy[§]

Abstract

Structured real canonical forms for matrices in $\mathbb{R}^{n\times n}$ that are symmetric or skew-symmetric about the anti-diagonal as well as the main diagonal are presented, and Jacobi algorithms for solving the complete eigenproblem for three of these four classes of matrices are developed. Based on the direct solution of 4×4 subproblems constructed via quaternions, the algorithms calculate structured orthogonal bases for the invariant subspaces of the associated matrix. In addition to preserving structure, these methods are inherently parallelizable, numerically stable, and show asymptotic quadratic convergence.

Key words. canonical form, eigenvalues, eigenvectors, Jacobi method, double structure preserving, symmetric, persymmetric, skew-symmetric, perskew-symmetric, centrosymmetric, perplectic, quaternion, tensor product, Lie algebra, Jordan algebra, bilinear form.

AMS subject classification. 65F15, 15A18, 15A21, 15A57, 15A69

1 Introduction

The numerical solution of structured eigenproblems is often called for in practical applications. In this paper we focus on four types of doubly structured real matrices — those that have symmetry or skew-symmetry about the anti-diagonal as well as the main diagonal. Instances where such matrices arise include the control of mechanical and electrical vibrations, where the eigenvalues and eigenvectors of Gram matrices that are symmetric about both diagonals play a fundamental role [23].

*Numerical Analysis Report No. 427, Manchester Center for Computational Mathematics, Manchester, England, May 2003. Revised February 2004.

†Department of Mathematics, University of Manchester, Manchester, M13 9PL, England (`smackey@ma.man.ac.uk`).

‡Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA (`nil.mackey@wmich.edu`, `http://homepages.wmich.edu/~mackey/`). This work was supported in part by NSF Grant CCR–9619514, and Engineering and Physical Sciences Research Council Visiting Fellowship GR/515563.

§Applied Mathematics and Scientific Computation Program, University of Maryland, College Park, MD USA (`ddunlavy@cs.umd.edu`, `http://www.math.umd.edu/~ddunlavy`).

In addition to presenting doubly structured real canonical forms for these four classes of matrices, we develop structure-preserving Jacobi algorithms to solve the eigenproblem for three of these classes. A noteworthy advantage of these methods is that the rich eigenstructure of the initial matrix is not obscured by rounding errors during the computation. Such algorithms also exhibit greater numerical stability, and are likely to be strongly backward stable [25]. Storage requirements are appreciably lowered by working with a truncated form of the matrix. Because our algorithms are Jacobi-like, they are readily adaptable for parallel implementation.

2 Automorphism groups, Lie and Jordan algebras

A number of important classes of real matrices can be profitably viewed as operators associated with a non-degenerate bilinear form $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$. (Complex bilinear or sesquilinear forms yield corresponding complex classes of matrices.)

$$\mathbb{G} = \{G \in \mathbb{R}^{n\times n} : \langle Gx, Gy\rangle = \langle x, y\rangle, \ \forall x, y \in \mathbb{R}^n\} \tag{2.1a}$$
$$\mathbb{L} = \{A \in \mathbb{R}^{n\times n} : \langle Ax, y\rangle = -\langle x, Ay\rangle, \ \forall x, y \in \mathbb{R}^n\} \tag{2.1b}$$
$$\mathbb{J} = \{A \in \mathbb{R}^{n\times n} : \langle Ax, y\rangle = \langle x, Ay\rangle, \ \forall x, y \in \mathbb{R}^n\} \tag{2.1c}$$

It follows that $\mathbb{G}$ is a multiplicative group, $\mathbb{L}$ is a subspace, closed under the Lie bracket defined by $[A, B] = AB - BA$, and $\mathbb{J}$ is a subspace closed under the Jordan product defined by $\{A, B\} = \frac{1}{2}(AB + BA)$. We will refer to $\mathbb{G}$, $\mathbb{L}$, and $\mathbb{J}$ as the automorphism group, Lie algebra and Jordan algebra, respectively, of the bilinear form $\langle \cdot, \cdot \rangle$. For our purposes, the most significant relationship between these three algebraic structures is that $\mathbb{L}$ and $\mathbb{J}$ are invariant under similarities by matrices in $\mathbb{G}$.

Proposition 2.1. *For any non-degenerate bilinear form on* $\mathbb{R}^n$*,*

$$A \in \mathbb{L},\ G \in \mathbb{G} \ \Rightarrow\ G^{-1}A\,G \in \mathbb{L}\,; \quad A \in \mathbb{J},\ G \in \mathbb{G} \ \Rightarrow\ G^{-1}A\,G \in \mathbb{J}.$$

Proof. Suppose $A \in \mathbb{L}$, $G \in \mathbb{G}$. Then for all $x, y \in \mathbb{R}^n$,

$$\begin{aligned}
\langle G^{-1}A\,Gx,\ y\rangle &= \langle G\,G^{-1}A\,Gx,\ Gy\rangle && \text{by (2.1a)}\\
&= \langle Gx,\ -A\,Gy\rangle && \text{by (2.1b)}\\
&= \langle G^{-1}Gx,\ -G^{-1}A\,Gy\rangle && \text{by (2.1a)}\\
&= \langle x,\ -G^{-1}A\,Gy\rangle
\end{aligned}$$

Thus $G^{-1}AG \in \mathbb{L}$. The second assertion is proved in a similar manner. □

Two familiar bilinear forms, $\langle x, y\rangle = x^T y$ and $\langle x, y\rangle = x^T J_{2p} y$ where $J_{2p} = \left[\begin{smallmatrix} 0 & I_p \\ -I_p & 0 \end{smallmatrix}\right]$, give rise to well-known $(\mathbb{G}, \mathbb{L}, \mathbb{J})$ triples, as noted in Table 2.1. Less familiar, perhaps, is the triple associated with the form $\langle x, y\rangle = x^T R_n\, y$ where R_n is the $n \times n$ matrix with 1's on the antidiagonal, and 0's elsewhere:

$$R_n \stackrel{\text{def}}{=} \begin{bmatrix} & & 1 \\ & ⋰ & \\ 1 & & \end{bmatrix}. \tag{2.2}$$

Letting $p\mathcal{S}(n)$ denote the Jordan algebra of this bilinear form, we see from (2.1c), that

$$p\mathcal{S}(n) = \{A \in \mathbb{R}^{n\times n} \ : \ A^T R_n = R_n A\} = \{A \in \mathbb{R}^{n\times n} \ : \ (R_n A)^T = R_n A\}\,. \tag{2.3}$$

It follows that matrices in $p\mathcal{S}(n)$ are symmetric about the anti-diagonal; they are often called the *persymmetric* matrices. Similarly, the Lie algebra consists of matrices that are skew-symmetric about the anti-diagonal,

$$p\mathcal{K}(n) = \{A \in \mathbb{R}^{n\times n} \ : \ A^T R_n = -R_n A\} = \{A \in \mathbb{R}^{n\times n} \ : \ (R_n A)^T = -R_n A\} \tag{2.4}$$

called, by analogy, the *perskew-symmetric* matrices. On the other hand, the automorphism group does not appear to have been specifically named. Yielding to whimsy, we will refer to this $\mathbb{G}$ as the *perplectic* group:

$$\mathrm{P}(n) = \{P \in \mathbb{R}^{n\times n} \ : \ P^T R_n P = R_n\}\,. \tag{2.5}$$

Note that $\mathrm{P}(n)$ is isomorphic as a group to the real pseudo-orthogonal group $\mathrm{O}(\lceil \frac{n}{2}\rceil, \lfloor \frac{n}{2}\rfloor)$, although the matrices in these two groups are quite different.

Bilinear Form $\langle x, y\rangle$	**Automorphism Group** $\mathbb{G}$	**Lie Algebra** $\mathbb{L}$	**Jordan Algebra** $\mathbb{J}$
$\langle x, y\rangle = x^T y$	Orthogonals	Skew-symmetrics	Symmetrics
$\langle x, y\rangle = x^T J_{2p} y$	Symplectics	Hamiltonians	Skew-Hamiltonians
$\langle x, y\rangle = x^T R_n\, y$	Perplectics	Perskew-symmetrics	Persymmetrics

Table 2.1: Examples of structured matrices associated with some bilinear forms

2.1 Flip operator

Following Reid [23] we define the "flip" operation $(\ \ )^F$, whose effect is to transpose a matrix across its anti-diagonal:

Definition 2.2. $\quad A^F \ := \ RA^T R$

One can verify that flipping is the adjoint with respect to the bilinear form $\langle x, y\rangle = x^T R_n\, y$; that is, for any $A \in \mathbb{R}^{n\times n}$ we have

$$\langle Ax, y\rangle = \langle x, A^F y\rangle, \quad \forall\, x, y \in \mathbb{R}^n. \tag{2.6}$$

Consequently the following properties of the flip operation are not surprising:

$$(B^F)^F = B, \qquad (AB)^F = B^F A^F, \qquad (B^F)^{-1} = (B^{-1})^F = B^{-F}. \tag{2.7}$$

It now follows immediately from (2.3), (2.4), and (2.5), or directly from (2.1) using the characterization of $(\cdot)^F$ as an adjoint, that

$$A \text{ is persymmetric } \Leftrightarrow A^F = A, \tag{2.8a}$$

$$A \text{ is perskew-symmetric } \Leftrightarrow A^F = -A, \tag{2.8b}$$

$$A \text{ is perplectic } \Leftrightarrow A^F = A^{-1}. \tag{2.8c}$$

The following proposition uses (2.8c) to determine when a $2n \times 2n$ block-upper-triangular matrix is perplectic.

Proposition 2.3. *Let $B, C, X \in \mathbb{R}^{n\times n}$. Then $\left[\begin{smallmatrix} B & X \\ 0 & C \end{smallmatrix}\right]$ is perplectic iff $C = B^{-F}$ and BX^F is perskew-symmetric.*

Proof. With $A = \left[\begin{smallmatrix} B & X \\ 0 & C \end{smallmatrix}\right]$, we have $A^F = \left[\begin{smallmatrix} C^F & X^F \\ 0 & B^F \end{smallmatrix}\right]$. Then $A^F = A^{-1}$ iff

$$AA^F = \begin{bmatrix} B & X \\ 0 & C \end{bmatrix} \begin{bmatrix} C^F & X^F \\ 0 & B^F \end{bmatrix} = \begin{bmatrix} BC^F & BX^F + XB^F \\ 0 & CB^F \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

B and C must be invertible, since any perplectic matrix is invertible. Equating corresponding blocks yields $C = B^{-F}$ and $BX^F = -XB^F = -(BX^F)^F$. $\square$

Analogously, one can show that $\left[\begin{smallmatrix} B & 0 \\ X & C \end{smallmatrix}\right]$ is perplectic iff $C = B^{-F}$ and $B^F X$ is perskew-symmetric. Interesting special cases include the block-diagonal perplectics, $\left[\begin{smallmatrix} B & 0 \\ 0 & C \end{smallmatrix}\right]$ with $C = B^{-F}$, and the perplectic shears, $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$ with X perskew-symmetric.[1]

The condition that BX^F be perskew-symmetric can also be expressed as

$$BX^F + XB^F = 0 \;\Leftrightarrow\; X^F B^{-F} = -B^{-1}X \;\Leftrightarrow\; (B^{-1}X)^F = -B^{-1}X,$$

that is, $B^{-1}X$ is perskew-symmetric. It is of interest to compare Proposition 2.3 with analogous results for symplectic block-upper-triangular matrices used in [9, 11]. There it is shown that

$$\left[\begin{smallmatrix} B & X \\ 0 & C \end{smallmatrix}\right] \text{ is symplectic } \;\Leftrightarrow\; C = B^{-T} \text{ and } B^{-1}X \text{ is symmetric},$$

with special cases the block-diagonal symplectics, $\left[\begin{smallmatrix} B & 0 \\ 0 & C \end{smallmatrix}\right]$ with $C = B^{-T}$, and the symplectic shears, $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$ with X symmetric. These concrete examples illustrate that, by contrast with the orthogonal groups, the perplectic and symplectic groups are not compact.

3 Perplectic orthogonals

Since orthogonal matrices are perfectly conditioned, and perplectic similarities preserve structure, perplectic orthogonal similarity transformations are ideal tools for the numerical solution of persymmetric and perskew-symmetric eigenproblems. From (2.5) it follows that the perplectic orthogonal group, which we denote by $\mathrm{PO}(n)$, is given by

$$\mathrm{PO}(n) = \{P \in \mathrm{O}(n) \mid R_n P = P R_n\}, \tag{3.1}$$

[1] It can be shown that every $2n \times 2n$ block-upper-triangular perplectic matrix $\left[\begin{smallmatrix} B & X \\ 0 & C \end{smallmatrix}\right]$ can be uniquely expressed as the product of a block-diagonal perplectic and a perplectic shear. The analogous factorization for block-upper-triangular symplectics was mentioned in [9], [11].

where $\mathrm{O}(n)$ is the $n \times n$ orthogonal group. Alternatively, one may characterize $\mathrm{PO}(n)$ as the set of all centrosymmetric orthogonal matrices.

Each perplectic orthogonal group $\mathrm{PO}(n)$ is a Lie group, so the dimension of $\mathrm{PO}(n)$ as a manifold is the same as the vector space dimension of its corresponding Lie algebra, the $n \times n$ skew-symmetric persymmetric matrices. These dimensions are recorded in Table 3.1 along with the dimensions of the full orthogonal groups for comparison. Note the 0-dimensionality of $\mathrm{PO}(2)$; this group contains only four elements, $\pm I_2$, $\pm R_2$.

n	2	3	4	5	$\dots$	n(even)	n (odd)
$\dim \mathrm{PO}(n)$	0	1	2	4	$\dots$	$\frac{1}{4}n(n-2)$	$\frac{1}{4}(n-1)^2$
$\dim \mathrm{O}(n)$	1	3	6	10	$\dots$	$\frac{1}{2}n(n-1)$	$\frac{1}{2}n(n-1)$

Table 3.1: Dimensions of $\mathrm{PO}(n)$ and $\mathrm{O}(n)$

Another basic property of $\mathrm{PO}(n)$ is its lack of connectedness. This contrasts with the symplectic orthogonal groups $\mathrm{SpO}(2n)$, which are always connected[2]. Since $\mathrm{PO}(n)$ is isomorphic to $\mathrm{O}(\lceil \frac{n}{2} \rceil) \times \mathrm{O}(\lfloor \frac{n}{2} \rfloor)$, it follows that it has four connected components. Concrete descriptions of these four components when $n = 3$, 4 are given in Appendix B.

The reason to raise the connectedness issue here is that our algorithms achieve their goals using only the matrices in $\mathrm{PO}_I(n)$, the connected component of $\mathrm{PO}(n)$ that contains the identity matrix I_n. This component is always a normal subgroup of $\mathrm{PO}(n)$ comprised only of rotations (orthogonal matrices U with $\det U = 1$). The exclusive use of $\mathrm{PO}_I(n)$ means "far-from-identity" transformations are avoided, which in turn promotes good convergence behavior of our algorithms.

4 Role of the quaternions

As has been pointed out in the case of real Hamiltonian and skew-Hamiltonian matrices [3], [10], a structure preserving Jacobi algorithm based on 2×2 subproblems is hampered by the fact that many of the off-diagonal elements are inaccessible to direct annihilation. For any 2×2 based Jacobi algorithm for persymmetric or perskew-symmetric matrices, the problem is even more acute: with $\mathrm{PO}(2) = \{\pm I_2, \pm R_2\}$, there are effectively no 2×2 structure preserving similarities with which to transform the matrix.

Following the strategy used in [10], [17], these difficulties can be overcome by using quaternions to construct simple closed form, real solutions to real doubly-structured 4×4 eigenproblems, and then building Jacobi algorithms for the corresponding $n \times n$ eigenproblems using these 4×4 solutions as a base.

The $n \times n$ skew-symmetric perskew-symmetric case, however, presents an additional challenge: when $n = 4$, such a matrix is already in canonical form, since no perplectic

[2]In [15] the group $\mathrm{SpO}(2n)$ is shown to be the continuous image of the complex unitary group $\mathrm{U}(n)$, which is known to be connected.

orthogonal similarity can reduce it further. A structure preserving Jacobi algorithm for these "doubly skewed" matrices must necessarily be based on the solution of larger subproblems, and this remains an open problem.

4.1 The quaternion tensor square $\mathbb{H} \otimes \mathbb{H}$

The connection between the quaternions

$$\mathbb{H} \;=\; \{q = q_0 + q_1 i + q_2 j + q_3 k \;:\; q_0, q_1, q_2, q_3 \in \mathbb{R}, \quad i^2 = j^2 = k^2 = ijk = -1\}$$

and 4×4 real matrices has been exploited before [10], [12], [17]. In particular, the algebra isomorphism between $\mathbb{R}^{4\times4}$ and the quaternion tensor $\mathbb{H}\otimes\mathbb{H}$ was used in [17] to show that real 4×4 symmetric and skew-symmetric matrices have a convenient quaternion characterization, and again in [10] to develop a quaternion representation for real 4×4 Hamiltonian and skew-Hamiltonian matrices. Since we will use this isomorphism to characterize real 4×4 persymmetric and perskew-symmetric matrices, a brief description of it is included here.

For each $(p, q) \in \mathbb{H} \times \mathbb{H}$, let $f(p, q) \in \mathbb{R}^{4\times4}$ denote the matrix representation of the real linear map on $\mathbb{H}$ defined by $v \mapsto pv\overline{q}$, using the standard basis $\{1, i, j, k\}$. Here $\overline{q}$ denotes the conjugate $q_0 - q_1 i - q_2 j - q_3 k$. The map $f : \mathbb{H} \times \mathbb{H} \to \mathbb{R}^{4\times4}$ is clearly bilinear, and consequently induces a unique *linear* map $\phi : \mathbb{H}\otimes\mathbb{H} \to \mathbb{R}^{4\times4}$ such that $\phi(p \otimes q) = f(p, q)$.

From the definition of ϕ it follows that

$$\phi(p\otimes 1) = \begin{bmatrix} p_0 & -p_1 & -p_2 & -p_3 \\ p_1 & p_0 & -p_3 & p_2 \\ p_2 & p_3 & p_0 & -p_1 \\ p_3 & -p_2 & p_1 & p_0 \end{bmatrix}, \quad \phi(1\otimes q) = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \\ -q_1 & q_0 & -q_3 & q_2 \\ -q_2 & q_3 & q_0 & -q_1 \\ -q_3 & -q_2 & q_1 & q_0 \end{bmatrix}. \tag{4.1}$$

It can be shown that ϕ is an isomorphism of algebras [2], [21]. The tensor multiplication rule $(a \otimes b)(a' \otimes b') = (aa' \otimes bb')$ then implies that the matrices in (4.1) commute, and their product is $\phi(p \otimes q)$. From (4.1) it also follows that

$$\phi(\overline{p} \otimes 1) = (\phi(p \otimes 1))^T, \quad \phi(1 \otimes \overline{q}) = (\phi(1 \otimes q))^T. \tag{4.2}$$

Since conjugation in $\mathbb{H}\otimes\mathbb{H}$ is determined by extending the rule $\overline{p \otimes q} = \overline{p} \otimes \overline{q}$ linearly to all of $\mathbb{H}\otimes\mathbb{H}$, we see that ϕ preserves more than the algebra structure: *conjugation in* $\mathbb{H}\otimes\mathbb{H}$ *corresponds, via* ϕ*, to transpose in* $\mathbb{R}^{4\times4}$.

By the usual abuse of notation, we will use $p \otimes q$ to stand for the matrix $\phi(p \otimes q)$, both to simplify notation and to emphasize the identification of $\mathbb{H}\otimes\mathbb{H}$ with $\mathbb{R}^{4\times4}$.

4.2 Rotations of $\mathbb{R}^3$ and $\mathbb{R}^4$

The correspondence between general rotations of $\mathbb{R}^3$ and $\mathbb{R}^4$ and the algebra of quaternions goes back to Hamilton and Cayley [4], [5], [13]. Briefly put in the language of section 4.1, every element of SO(4) can be expressed as $x \otimes y$, where x and y are quaternions of unit length. This means that the map $q \mapsto xq\overline{y}$ can be interpreted as a rotation of $\mathbb{R}^4$. Similarly, every element of SO(3) can be realized as $x \otimes x$ for some unit quaternion x. In this case the map $q \mapsto xq\overline{x}$ keeps the real part of q invariant, and can be interpreted as a rotation acting on the subspace of pure quaternions, $\mathbb{P} = \{p_1 i + p_2 j + p_3 k \ : \ p_1, p_2, p_3 \in \mathbb{R}\} \cong \mathbb{R}^3$.

There is a useful and direct relation between the coordinates of a unit quaternion $x = x_0 + x_1 i + x_2 j + x_3 k$ and the geometry of the associated rotation $x \otimes x \in \mathrm{SO}(3)$.

Proposition 4.1. *Let x be a unit quaternion. Then $x \otimes x \in \mathrm{SO}(3)$ is a rotation with axis along the vector given by the pure quaternion part, (x_1, x_2, x_3), and angle θ determined by the real part, $x_0 = \cos(\theta/2)$.*

Proof. See, for example, [6], [22]. $\square$

The following proposition, adapted from [12] and used in [10], will be useful in section 5.

Proposition 4.2. *Suppose $a, b \in \mathbb{P}$ are nonzero pure quaternions such that $|ba| - ba \neq 0$ (equivalently, such that $a/|a| \neq -b/|b|$), and let x be the unit quaternion*

$$x \ = \ \frac{|ba| - ba}{|\,|ba| - ba\,|} \ = \ \frac{|b|\,|a| - ba}{|\ |b|\,|a| - ba\ |} \ . \tag{4.3}$$

Then $x \otimes x \in \mathrm{SO}(3)$ rotates a into alignment with b. Furthermore, if a and b are linearly independent, and x is chosen as in (4.3), then $x \otimes x$ is the smallest angle rotation that sends a into alignment with b.

4.3 4×4 perplectic rotations

Let $P \in \mathrm{SO}(4)$. Then P can be expressed as $x \otimes y$ where x, y are unit quaternions. If P is also perplectic, then by (3.1), P commutes with $R_4 = j \otimes i$. Hence

$$\begin{aligned} P \in \mathrm{P}(4) \cap \mathrm{SO}(4) \ &\Leftrightarrow \ (x \otimes y)(j \otimes i) = (j \otimes i)(x \otimes y) \\ &\Leftrightarrow \ xj \otimes yi = jx \otimes iy \\ &\Leftrightarrow \ (xj = jx \ \text{ and } \ yi = iy) \ \text{ or } \ (xj = -jx \ \text{ and } \ yi = -iy). \end{aligned}$$

The first alternative implies $x \in \mathrm{span}\,\{1, j\}$ and $y \in \mathrm{span}\{1, i\}$, while the second implies $x \in \mathrm{span}\{i, k\}$ and $y \in \mathrm{span}\{j, k\}$. These two alternatives correspond to the two connected components of 4×4 perplectic rotations, with the first alternative describing $\mathrm{PO}_I(4)$, the connected component containing the identity. This quaternion parametrization

$$\mathrm{PO}_I(4) \ = \ \{x \otimes y \ : \ |x| = |y| = 1, \ x \in \mathrm{span}\{1, j\}, \ y \in \mathrm{span}\{1, i\}\}\,, \tag{4.4}$$

together with the geometric characterization given in the following proposition will be used to construct structure preserving transformations for the algorithms in this paper.

Proposition 4.3. *Let x, y be unit quaternions such that $x \otimes y \in \mathrm{PO}_I(4)$. Then the axes of the 3-dimensional rotations $x \otimes x$ and $y \otimes y$ lie along $j = (0, 1, 0)$ and $i = (1, 0, 0)$ respectively.*

Proof. When $x \otimes y \in \mathrm{PO}_I(4)$, Proposition 4.1 together with (4.4) imply that the angles of both rotations can be freely chosen, but their axes must lie along j and i, respectively. $\square$

4.4 Similarities by rotations

By using quaternions, the computation of rotational similarities in $\mathbb{R}^{4\times4}$ becomes tractable. This was used to advantage in [10], [17], and will once again be exploited here.

Let $a, b \in \mathbb{H}$ be given. If x, y are unit quaternions, then the product $(x \otimes y)(a \otimes b)(\overline{x} \otimes \overline{y}) \in \mathbb{H} \otimes \mathbb{H}$ represents a similarity transformation on $\phi(a \otimes b) \in \mathbb{R}^{4\times4}$ by $\phi(x \otimes y) \in \mathrm{SO}(4)$. On the other hand,

$$(x \otimes y)(a \otimes b)(\overline{x} \otimes \overline{y}) = (xa\overline{x}) \otimes (yb\overline{y}). \tag{4.5}$$

By Section 4.2, this means that the pure quaternion part of a is rotated by the 3-dimensional rotation $x \otimes x$, while an independent rotation, $y \otimes y \in \mathrm{SO}(3)$ rotates the pure quaternion part of b. Since every element of $\mathbb{H} \otimes \mathbb{H}$ is a real linear combination of terms of the form $a \otimes b$, the effect of a similarity by $x \otimes y \in \mathrm{SO}(4)$ can be reduced to the action of a pair of independent 3-dimensional rotations.

4.5 Simultaneous splittings

When viewed in $\mathbb{R}^{4\times4}$ via the isomorphism ϕ, the standard basis $\mathcal{B} = \{1 \otimes 1, 1 \otimes i, \ldots, k \otimes j, k \otimes k\}$ of $\mathbb{H} \otimes \mathbb{H}$ was shown in [10], [17], to consist of matrices that are symmetric or skew-symmetric as well as Hamiltonian or skew-Hamiltonian. Something even more remarkable is true. Each of these sixteen matrices is also either persymmetric or perskew-symmetric. Thus the quaternion basis simultaneously exhibits no less than three direct sum decompositions of $\mathbb{R}^{4\times4}$ into $\mathbb{J} \oplus \mathbb{L}$:

$$\begin{array}{rcl} \{\text{Symmetrics}\} & \oplus & \{\text{Skew-symmetrics}\} \\ \{\text{Skew-Hamiltonians}\} & \oplus & \{\text{Hamiltonians}\} \\ \{\text{Persymmetrics}\} & \oplus & \{\text{Perskew-symmetrics}\} \end{array}$$

This is shown in Tables 4.1-4.3. For the matrix representation of the quaternion basis, see Appendix A.

An elegant explanation for why $\mathcal{B}$ has this simultaneous splitting property can be outlined as follows:

- The correspondence between conjugation and transpose explains why each basis element is either symmetric or skew-symmetric. For example, $k \otimes j$ is its own conjugate, so the matrix $\phi(k \otimes j)$ must be symmetric.

$\otimes$	1	i	j	k
1	S	K	K	K
i	K	S	S	S
j	K	S	S	S
k	K	S	S	S

S = Symmetric
K = Skewsymmetric

Table 4.1:

$\otimes$	1	i	j	k
1	W	W	H	W
i	H	H	W	H
j	H	H	W	H
k	H	H	W	H

W = Skew–Hamiltonian
H = Hamiltonian

Table 4.2:

$\otimes$	1	i	j	k
1	pS	pK	pS	pS
i	pS	pK	pS	pS
j	pK	pS	pK	pK
k	pS	pK	pS	pS

pS = Persymmetric
pK = Perskew–symmetric

Table 4.3:

- Premultiplication by J_{2n}, the matrix that gives rise to the symplectic bilinear form, is a bijection that turns symmetric matrices into Hamiltonian ones and skew-symmetric matrices into skew-Hamiltonian ones. Similarly, the bijection given by premultiplication by R_n, the matrix associated with the perplectic bilinear form, turns symmetric matrices into persymmetric matrices and skew-symmetric matrices into perskew-symmetric ones.

- Up to sign, $\mathcal{B}$ is closed under multiplication. This is trivial to verify in $\mathbb{H} \otimes \mathbb{H}$. Now by a fortuitous concordance, both J_4 and R_4 belong to $\mathcal{B}$, since $J_4 = 1 \otimes j$, and $R_4 = j \otimes i$. Hence the effect of premultiplication by R_4 or J_4 is merely to permute (up to sign) the elements of $\mathcal{B}$. For example, since $k \otimes j$ is symmetric, and $R_4(k \otimes j) = (j \otimes i)(k \otimes j) = jk \otimes ij = i \otimes k$, it follows that $i \otimes k$ is persymmetric.

Thus one of the reasons why all three families of structures are simultaneously reflected in $\mathcal{B}$ is that the matrices I_4, J_4 and R_4 that define the underlying bilinear forms are themselves elements of $\mathcal{B}$. This suggests the possibility of further extensions: each of the sixteen quaternion basis elements could be used to define a non-degenerate bilinear form on $\mathbb{R}^4$, thus giving rise to sixteen $(\mathbb{G}, \mathbb{L}, \mathbb{J})$ triples on $\mathbb{R}^{4\times 4}$, which might then be extended in some way to triples of structured $n \times n$ matrices. However, these sixteen bilinear forms on $\mathbb{R}^4$ are not all distinct. In fact, they fall into exactly three equivalence classes. The bilinear form defined by I_4 is in a class by itself. The other nine symmetric matrices in $\mathcal{B}$ give rise to bilinear forms that are all equivalent to $\langle x, y\rangle = x^T R_4\, y$. The remaining six skew-symmetric matrices in $\mathcal{B}$ define forms that are each equivalent to $\langle x, y\rangle = x^T J_4\, y$. Thus the three $(\mathbb{G}, \mathbb{L}, \mathbb{J})$ triples defined in Table 2.1 are essentially the only ones with quaternion ties.

4.6 Quaternion dictionary

Using Tables 4.1 and 4.3, quaternion representations of structured classes of matrices relevant to this work can be constructed; these are listed in Table 4.4. For easy reference, the representation for rotations and perplectic rotations developed in sections 4.2 and 4.3 are also included in the table. For representations of symmetric or skew-symmetric Hamiltonian and skew-Hamiltonian matrices, the interested reader is referred to [10].

Table 4.4: Quaternion dictionary for some structured 4×4 matrices

	$\alpha,\,\beta,\,\gamma,\,\delta \in \mathbb{R}, \quad p,\,q,\,r \in \mathbb{P}$
Diagonal	$\alpha\,(1 \otimes 1) \,+\, \beta\,(i \otimes i) \,+\, \gamma\,(j \otimes j) \,+\, \delta\,(k \otimes k)$
Symmetric	$\alpha\,(1 \otimes 1) \,+\, p \otimes i \,+\, q \otimes j \,+\, r \otimes k$
Skew-symmetric	$p \otimes 1 \,+\, 1 \otimes q$

	$\alpha,\,\beta \in \mathbb{R}, \quad p,\,q,\,r \in \operatorname{span}\{i,k\},\ s \in \operatorname{span}\{j,k\}$
Persymmetric	$\alpha\,(1 \otimes 1) \,+\, \beta\,(j \otimes i) \,+\, p \otimes j \,+\, q \otimes k \,+\, r \otimes 1 \,+\, 1 \otimes s$
Symmetric persymmetric	$\alpha\,(1 \otimes 1) \,+\, \beta\,(j \otimes i) \,+\, p \otimes j \,+\, q \otimes k$
Skew-symmetric persymmetric	$r \otimes 1 \,+\, 1 \otimes s$
Perskew-symmetric	$r \otimes i \,+\, j \otimes s \,+\, \alpha\,(1 \otimes i) \,+\, \beta\,(j \otimes 1)$
Symmetric perskew-symmetric	$r \otimes i \,+\, j \otimes s$
Skew-symmetric perskew-symmetric	$\alpha\,(1 \otimes i) \,+\, \beta\,(j \otimes 1)$

	$\lvert x\rvert = \lvert y\rvert = 1,\ x,\,y \in \mathbb{H}$
Rotation	$x \otimes y$
Perplectic rotation	$x \otimes y, \quad x \in \operatorname{span}\{1,j\},\ y \in \operatorname{span}\{1,i\}$, OR $x \in \operatorname{span}\{i,k\},\ y \in \operatorname{span}\{j,k\}$

We now specify the quaternion parameters for each of the six types of structured 4×4 matrices listed in the second group of Table 4.4. This is done in terms of the matrix entries by using the matrix form of the basis $\mathcal{B}$ given in Appendix A.

If $A = [a_{\ell m}] = \alpha\,(1 \otimes 1) \,+\, \beta\,(j \otimes i) \,+\, p \otimes j \,+\, q \otimes k \,+\, r \otimes 1 \,+\, 1 \otimes s$ is a 4×4 real persymmetric matrix, then the scalars α, $\beta \in \mathbb{R}$, and the pure quaternion parameters p, q, $r \in \operatorname{span}\{i,k\}$, $s \in \operatorname{span}\{j,k\}$, are given by

$$\alpha = \tfrac{1}{2}(a_{11} + a_{22}) \tag{4.6a}$$
$$\beta = \tfrac{1}{4}\,(a_{14} + a_{23} + a_{32} + a_{41}) \tag{4.6b}$$
$$p = [p_1, p_2, p_3] = [\,\tfrac{1}{4}(-a_{14} + a_{23} + a_{32} - a_{41}),\ 0,\ \tfrac{1}{2}(a_{21} + a_{12})\,] \tag{4.6c}$$
$$q = [q_1, q_2, q_3] = [\,\tfrac{1}{2}(a_{13} + a_{31}),\ 0,\ \tfrac{1}{2}(a_{11} - a_{22})\,] \tag{4.6d}$$
$$r = [r_1, r_2, r_3] = [\,\tfrac{1}{2}(a_{21} - a_{12}),\ 0,\ \tfrac{1}{4}(-a_{14} - a_{23} + a_{32} + a_{41})\,] \tag{4.6e}$$
$$s = [s_1, s_2, s_3] = [\,0,\ \tfrac{1}{2}(a_{13} - a_{31}),\ \tfrac{1}{4}(a_{14} - a_{23} + a_{32} - a_{41})\,]. \tag{4.6f}$$

The corresponding calculation for a 4×4 real perskew-symmetric matrix $A = [a_{\ell m}] = r \otimes i \,+\, j \otimes s \,+\, \alpha\,(1 \otimes i) \,+\, \beta\,(j \otimes 1)$ yields even simpler equations for the scalars α, $\beta \in \mathbb{R}$ and the pure quaternions $r \in \operatorname{span}\{i,k\}$, $s \in \operatorname{span}\{j,k\}$.

$$\alpha = \tfrac{1}{2}(a_{12} - a_{21}) \tag{4.7a}$$
$$\beta = \tfrac{1}{2}(-a_{13} + a_{31}) \tag{4.7b}$$
$$r = [r_1, r_2, r_3] = [\,\tfrac{1}{2}(a_{11} + a_{22}),\ 0,\ -\tfrac{1}{2}(a_{13} + a_{31})\,] \tag{4.7c}$$
$$s = [s_1, s_2, s_3] = [\,0,\ \tfrac{1}{2}(a_{11} - a_{22}),\ -\tfrac{1}{2}(a_{12} + a_{21})\,]. \tag{4.7d}$$

Next, the four doubly structured classes are handled by specializing (4.6) – (4.7).

Type A: *Symmetric Persymmetric*

$$\alpha = \tfrac{1}{2}(a_{11} + a_{22}) \tag{4.8a}$$
$$\beta = \tfrac{1}{2}(a_{14} + a_{23}) \tag{4.8b}$$
$$p = [p_1, p_2, p_3] = [\,\tfrac{1}{2}(-a_{14} + a_{23}),\ 0,\ a_{12}\,] \tag{4.8c}$$
$$q = [q_1, q_2, q_3] = [\,a_{13},\ 0,\ \tfrac{1}{2}(a_{11} - a_{22})\,]. \tag{4.8d}$$

Type B: *Skew-symmetric Persymmetric*

$$r = [r_1, r_2, r_3] = [\,-a_{12},\ 0,\ -\tfrac{1}{2}(a_{14} + a_{23})\,] \tag{4.9a}$$
$$s = [s_1, s_2, s_3] = [\,0,\ a_{13},\ \tfrac{1}{2}(a_{14} - a_{23})\,]. \tag{4.9b}$$

Type C: *Symmetric perskew-symmetric*

$$r = [r_1, r_2, r_3] = [\,\tfrac{1}{2}(a_{11} + a_{22}),\ 0,\ -a_{13}\,] \tag{4.10a}$$
$$s = [s_1, s_2, s_3] = [\,0,\ \tfrac{1}{2}(a_{11} - a_{22}),\ -a_{12}\,]. \tag{4.10b}$$

Type D: *Skew-symmetric perskew-symmetric:*

$$\alpha = a_{12} \qquad\qquad \beta = a_{13} \tag{4.11}$$

5 Doubly structured 4×4 eigenproblems

Canonical forms via structure preserving similarities are now developed in closed form for 4×4 matrices of Type A, B, and C. This is done by reinterpreting these questions inside $\mathbb{H} \otimes \mathbb{H}$ as 3-dimensional geometric problems.

For a matrix A of Type D, it can be shown that no 4×4 perplectic orthogonal similarity can reduce A to a more condensed form. Indeed if one uses $W \in \mathrm{PO}_I(4)$, then $WAW^T = A$. This can be seen by using (4.5) with $a \otimes b$ replaced by the quaternion representation of a Type D matrix as given in Table 4.4:

$$(x \otimes y)\big(\alpha(1 \otimes i) + \beta(j \otimes 1)\big)(\overline{x} \otimes \overline{y}) = \alpha(1 \otimes yi\overline{y}) + \beta(xj\overline{x} \otimes 1) = \alpha(1 \otimes i) + \beta(j \otimes 1). \tag{5.1}$$

The last equality in (5.1) follows from (4.4). Other similarities from PO(4) can change A, but only in trivial ways: interchanging the roles of α, β, or changing their signs. Consequently a Jacobi algorithm for $n \times n$ skew-symmetric perskew-symmetric matrices cannot be based on 4×4 structured subproblems. Larger subproblems would need to be solved; finding closed form structure-preserving solutions for these remains under investigation.

5.1 4×4 symmetric persymmetric

Given a symmetric persymmetric matrix $A = \alpha(1 \otimes 1) + \beta(j \otimes i) + p \otimes j + q \otimes k \in \mathbb{R}^{4\times 4}$, to what extent can A be reduced to a simpler form by the similarity WAW^T

when $W = x \otimes y \in \mathrm{PO}(4)$? It is clear that the term $\alpha(1 \otimes 1)$ is invariant under all similarities. Converting the second term to matrix form yields $\beta(j \otimes i) = \beta R_4$. Since every $W \in \mathrm{PO}(4)$ commutes with R_4, the second term will also remain unaffected. Thus the reduced form of A will in general have terms on the main diagonal as well as the antidiagonal, and we conclude that A may be reduced, at best, to an "X-form" that will inherit the double symmetry of A:

$$\begin{bmatrix} \alpha_1 & & & \beta_1 \\ & \alpha_2 & \beta_2 & \\ & \beta_2 & \alpha_2 & \\ \beta_1 & & & \alpha_1 \end{bmatrix} \tag{5.2}$$

A matrix in this form will have eigenvalues given by $\alpha_1 \pm \beta_1$ and $\alpha_2 \pm \beta_2$. Now for the purpose of calculating a W that reduces A to X-form, we may assume without loss of generality that $A = p \otimes j + q \otimes k$. Thus we have

$$WAW^T = (xp\overline{x} \otimes yj\overline{y}) + (xq\overline{x} \otimes yk\overline{y}).$$

Recall from Table 4.4 that p, $q \in \mathrm{span}\{i, k\}$. The X-form of (5.2) would be achieved by taking $y = 1$ and rotating the pure quaternions p and q to multiples of i and k, respectively. But p and q are affected only by the rotation $x \otimes x$, which in general can align either p with $\pm i$, or q with $\pm k$, but not both. To overcome this difficulty we modify a strategy first used in [17] for general symmetric matrices.

Define a vector space isomorphism $\psi : \mathbb{P} \otimes \mathbb{P} \to \mathbb{R}^{3\times 3}$ as the unique linear extension of the map that sends $a \otimes b$ to the rank one matrix $ab^T \in \mathbb{R}^{3\times 3}$. Then we get

$$\begin{aligned} \psi(A) &= pe_2^T + qe_3^T \\ &= \begin{bmatrix} 0 & p_1 & q_1 \\ 0 & 0 & 0 \\ 0 & p_3 & q_3 \end{bmatrix} \\ &= \sigma_1 \begin{bmatrix} u_{11} \\ 0 \\ u_{21} \end{bmatrix} \begin{bmatrix} 0 \\ v_{11} \\ v_{21} \end{bmatrix}^T + \sigma_2 \begin{bmatrix} u_{12} \\ 0 \\ u_{22} \end{bmatrix} \begin{bmatrix} 0 \\ v_{12} \\ v_{22} \end{bmatrix}^T \qquad \text{by SVD} \qquad (5.3) \\ &= \psi(\sigma_1 u_1 \otimes v_1 + \sigma_2 u_2 \otimes v_2), \end{aligned}$$

where $\begin{bmatrix} u_{1i} & u_{2i} \end{bmatrix}^T$ and $\begin{bmatrix} v_{1i} & v_{2i} \end{bmatrix}^T$, $i = 1, 2$, are respectively the left and right singular vectors corresponding to the singular values $\sigma_1 \geq \sigma_2 \geq 0$ of $\left(\begin{smallmatrix} p_1 & q_1 \\ p_3 & q_3 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$.

Since ψ is an isomorphism, we have $A = p \otimes j + q \otimes k = \sigma_1(u_1 \otimes v_1) + \sigma_2(u_2 \otimes v_2)$. Because u_1, u_2 are orthogonal and lie in the i-k plane, a 3-dimensional rotation $x \otimes x$ with axis along j that aligns u_1 with k must also align u_2 with $\pm i$. Similarly, since v_1, v_2 are orthogonal vectors in the j-k plane, a rotation $y \otimes y$ with axis along i that aligns v_1 with k will align v_2 with $\pm j$. By Proposition 4.1, the unit quaternions x, y must lie in $\mathrm{span}\{1, j\}$ and $\mathrm{span}\{1, i\}$ respectively. Then $W = (x \otimes 1)(1 \otimes y) = x \otimes y$ will be in $\mathrm{PO}_I(4)$ by (4.4), and

$$WAW^T = \sigma_1(xu_1\overline{x} \otimes yv_1\overline{y}) + \sigma_2(xu_2\overline{x} \otimes yv_2\overline{y}) = \sigma_1(k \otimes k) \pm \sigma_2(i \otimes j)$$

is in X-form. Furthermore, since u_1 and v_1 are the singular vectors corresponding to the largest singular value σ_1, most of the "weight" of A has been sent to the main diagonal (represented here by $k \otimes k$), while the anti-diagonal (represented here by $i \otimes j$) carries the "secondary" weight.

An X-form can also be achieved by choosing $x \otimes y$ so that u_1 is aligned with i, and v_1 with j, effectively reversing the roles of the main diagonal and the anti-diagonal.

To calculate the unit quaternion x, use (4.3) with $a = u_1$, $b = k$; the computation of y is similar, this time with $a = v_1$, and $b = k$. The matrix forms of $x \otimes 1$ and $1 \otimes y$ can then be computed from (4.1); the product of these two commuting matrices yields W. Observe that to determine W, it suffices to find just one singular vector pair u_1, v_1, of a 2×2 matrix. In practise, one calculates the singular vectors corresponding to the largest singular value.

The computation of W involves the terms $\gamma = 1 + u_{21}$ and $\delta = 1 + v_{21}$. Thus subtractive cancellation can occur when u_{21} or v_{21} is negative, that is, when $u_1 = u_{11}i + u_{21}k$ or $v_1 = v_{11}j + v_{21}k$ require rotations by obtuse angles to bring them into alignment with k. By replacing u_1 by $-u_1$ and/or v_1 by $-v_1$ as needed, cancellation can be avoided, and the rotation angles will now be less than 90°(see Proposition 4.2). The computation of W is given in the following algorithm, which has been arranged for clarity, rather than optimality.

Algorithm 1 **(4×4 symmetric persymmetric).** *Given a symmetric persymmetric matrix $A \in \mathbb{R}^{4\times4}$, this algorithm computes a real perplectic orthogonal matrix $W \in \mathrm{PO}_I(4)$ such that WAW^T is in X-form as in* (5.2).

$p = \begin{bmatrix} \frac{1}{2}(a_{23} - a_{14}) & a_{12} \end{bmatrix}^T$ % from (4.8c)

$q = \begin{bmatrix} a_{13} & \frac{1}{2}(a_{11} - a_{22}) \end{bmatrix}^T$ % from (4.8d)

$\begin{bmatrix} U & \Sigma & V \end{bmatrix} := \mathtt{svd}\left(\begin{bmatrix} p & q \end{bmatrix}\right)$

$u = \begin{bmatrix} u_{11} & u_{21} \end{bmatrix}$ % $u_1 = u_{11}i + u_{21}k$, as in (5.3)

$v = \begin{bmatrix} v_{11} & v_{21} \end{bmatrix}$ % $v_1 = v_{11}j + v_{21}k$ as in (5.3)

% Change sign to avoid cancellation in computation of α, β

if $u_{21} < 0$ then $u = -u$ endif

if $v_{21} < 0$ then $v = -v$ endif

$\alpha = 1 + u_{21}$; $\beta = 1 + v_{21}$

$\gamma = \sqrt{2\alpha}$; $\delta = \sqrt{2\beta}$

$$W_x = \frac{1}{\gamma}\begin{bmatrix} \alpha & 0 & u_{11} & 0 \\ 0 & \alpha & 0 & -u_{11} \\ -u_{11} & 0 & \alpha & 0 \\ 0 & u_{11} & 0 & \alpha \end{bmatrix} \qquad \% \; W_x = x \otimes 1$$

$$W_y = \frac{1}{\delta}\begin{bmatrix} \beta & v_{11} & 0 & 0 \\ -v_{11} & \beta & 0 & 0 \\ 0 & 0 & \beta & -v_{11} \\ 0 & 0 & v_{11} & \beta \end{bmatrix} \qquad \% \; W_y = 1 \otimes y$$

$W = W_x W_y$ % WAW^T is now in X-form

5.2 4×4 skew-symmetric persymmetric

A skew-symmetric persymmetric matrix in $\mathbb{R}^{4\times4}$ is of the form $A = r \otimes 1 + 1 \otimes s$ where $r \in \operatorname{span}\{i, k\}$ and $s \in \operatorname{span}\{j, k\}$. Consequently one can choose a rotation $x \otimes x$ with axis along j that aligns r with k, and an independent rotation $y \otimes y$ with axis along i that aligns s with k. Then $W = x \otimes y \in \mathrm{PO}_I(4)$ by (4.4), and

$$\begin{aligned} WAW^T &= xr\overline{x} \otimes 1 + 1 \otimes ys\overline{y} \\ &= |r|k \otimes 1 + |s|1 \otimes k \\ &= \begin{bmatrix} 0 & 0 & 0 & |s| - |r| \\ 0 & 0 & -|s| - |r| & 0 \\ 0 & |s| + |r| & 0 & 0 \\ -|s| + |r| & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{5.4}$$

To calculate the unit quaternion x, use (4.3) with $a = r$, $b = k$; the computation of y is similar, this time with $a = s$, and $b = k$. The matrix forms of $x \otimes 1$ and $1 \otimes y$ can then be computed from (4.1); the product of these two commuting matrices yields W.

The computation of W involves the terms $\alpha = \|r\|_2 + r_2$ and $\beta = \|s\|_2 + s_2$. Thus subtractive cancellation can occur when r_2 or s_2 is negative, that is, when $r = r_1 i + r_2 k$, or $s = s_1 j + s_2 k$ require rotations by obtuse angles to bring them into alignment with k. By replacing r by $-r$ and/or s by $-s$ as needed, cancellation can be avoided, and the rotation angles will now be less than $90°$(see Proposition 4.2). The computation of W is given in the following algorithm, which has been arranged for clarity, rather than optimality.

Algorithm 2 (4×4 **skew-symmetric persymmetric).** *Given a skew-symmetric persymmetric matrix $A \in \mathbb{R}^{4\times4}$, this algorithm computes a real perplectic orthogonal matrix $W \in \mathrm{PO}_I(4)$ such that WAW^T is in anti-diagonal canonical form as in* (5.4).

$r = \begin{bmatrix} -a_{12} & -\frac{1}{2}(a_{14} + a_{23}) \end{bmatrix}$ % from (4.9a)

$s = \begin{bmatrix} a_{13} & \frac{1}{2}(a_{14} - a_{23}) \end{bmatrix}$ % from (4.9b)

% Change sign to avoid cancellation in computation of α, β

if $r_2 < 0$ then $r = -r$ endif

if $s_2 < 0$ then $s = -s$ endif

$\alpha = \|r\|_2 + r_2$; $\beta = \|s\|_2 + s_2$

$\gamma = \| \begin{bmatrix} r_1 & \alpha \end{bmatrix} \|_2$; $\delta = \| \begin{bmatrix} s_1 & \beta \end{bmatrix} \|_2$

$$W_x = \frac{1}{\gamma} \begin{bmatrix} \alpha & 0 & r_1 & 0 \\ 0 & \alpha & 0 & -r_1 \\ -r_1 & 0 & \alpha & 0 \\ 0 & r_1 & 0 & \alpha \end{bmatrix} \quad \% \, W_x = x \otimes 1$$

$$W_y = \frac{1}{\delta} \begin{bmatrix} \beta & s_1 & 0 & 0 \\ -s_1 & \beta & 0 & 0 \\ 0 & 0 & \beta & -s_1 \\ 0 & 0 & s_1 & \beta \end{bmatrix} \quad \% \, W_y = 1 \otimes y$$

$W = W_x W_y$ % WAW^T is now in canonical form as in (5.4)

5.3 4×4 symmetric perskew-symmetric

If $A \in \mathbb{R}^{4\times4}$ is symmetric perskew-symmetric, then $A = r \otimes i + j \otimes s$ where $r \in \operatorname{span}\{i,k\}$ and $s \in \operatorname{span}\{j,k\}$. Choose a unit quaternion x so that the rotation $x \otimes x$ has axis along j, and $xr\overline{x}$ is a multiple of i. Similarly choose a rotation $y \otimes y$ with axis along i that sends s to a multiple of j. Setting $W = x \otimes y$, we see from (4.4) that $W \in \mathrm{PO}_I(4)$, and

$$\begin{aligned} WAW^T &= xr\overline{x} \otimes i + j \otimes ys\overline{y} \;=\; |r| i \otimes i + |s| j \otimes j \\ &= \begin{bmatrix} |r|+|s| & 0 & 0 & 0 \\ 0 & |r|-|s| & 0 & 0 \\ 0 & 0 & -|r|+|s| & 0 \\ 0 & 0 & 0 & -|r|-|s| \end{bmatrix}. \end{aligned} \tag{5.5}$$

To calculate the unit quaternion x, use (4.3) with $a = r$, $b = i$; the computation of y is similar, this time with $a = s$, and $b = j$. The matrix forms of $x \otimes 1$ and $1 \otimes y$ can then be computed from (4.1); the product of these two commuting matrices yields W.

The computation of W involves the terms $\alpha = \|r\|_2 + r_1$ and $\beta = \|s\|_2 + s_1$. Thus subtractive cancellation can occur when r_1 or s_1 is negative, that is, when $r = r_1 i + r_2 k$, or $s = s_1 j + s_2 k$ require rotations by obtuse angles to bring them into alignment with i, j, respectively. By replacing r by $-r$ and/or s by $-s$ as needed, cancellation can be avoided, and the rotation angles will now be less than $90°$(see Proposition 4.2). The computation of W is given in the following algorithm, which has been arranged for clarity, rather than optimality.

Algorithm 3 (4×4 symmetric perskew-symmetric). *Given a symmetric perskew-symmetric matrix $A \in \mathbb{R}^{4\times4}$, this algorithm computes a real perplectic orthogonal matrix $W \in \mathrm{PO}_I(4)$ such that WAW^T is in diagonal canonical form as in* (5.5).

$r = \begin{bmatrix} \frac{1}{2}(a_{11}+a_{22}) & -a_{13} \end{bmatrix}$ % from (4.10a)
$s = \begin{bmatrix} \frac{1}{2}(a_{11}-a_{22}) & -a_{12} \end{bmatrix}$ % from (4.10b)
% Change sign to avoid cancellation in computation of α, β
if $r_1 < 0$ then $r = -r$ endif
if $s_1 < 0$ then $s = -s$ endif
$\alpha = \|r\|_2 + r_1$; $\beta = \|s\|_2 + s_1$
$\gamma = \| \begin{bmatrix} \alpha & r_2 \end{bmatrix} \|_2$; $\delta = \| \begin{bmatrix} \beta & s_2 \end{bmatrix} \|_2$

$$W_x = \frac{1}{\gamma} \begin{bmatrix} \alpha & 0 & -r_2 & 0 \\ 0 & \alpha & 0 & r_2 \\ r_2 & 0 & \alpha & 0 \\ 0 & -r_2 & 0 & \alpha \end{bmatrix} \quad \% \; W_x = x \otimes 1$$

$$W_y = \frac{1}{\delta} \begin{bmatrix} \beta & -s_2 & 0 & 0 \\ s_2 & \beta & 0 & 0 \\ 0 & 0 & \beta & s_2 \\ 0 & 0 & -s_2 & \beta \end{bmatrix} \quad \% \; W_y = 1 \otimes y$$

$W = W_x W_y$ % WAW^T is now in canonical form as in (5.5)

6 Doubly structured 3×3 eigenproblems

As we shall see in section 8, when n is odd, Jacobi algorithms for $n \times n$ matrices in the classes considered in this paper also require the solution to 3×3 eigenproblems.

6.1 PO(3)

Rather than working via the quaternion characterization of SO(3), a useful parametrization of PO(3) that exhibits its four connected components can be obtained directly from (3.1). Two of these components form the intersection of PO(3) with the group of rotations SO(3). Our algorithms will only use matrices from $\mathrm{PO}_I(3)$, the connected component containing the identity, given by

$$\mathrm{PO}_I(3) = \left\{ W(\theta) = \frac{1}{2} \begin{bmatrix} c+1 & \sqrt{2}s & c-1 \\ -\sqrt{2}s & 2c & -\sqrt{2}s \\ c-1 & \sqrt{2}s & c+1 \end{bmatrix} : c = \cos\theta, s = \sin\theta, \theta \in [0, 2\pi) \right\}. \tag{6.1}$$

This restriction to $\mathrm{PO}_I(3)$ ensures, just as in the 4×4 case, that "far-from-identity" rotations are avoided in our algorithms. Details of the derivation of (6.1) as well as parametrizations for the other three connected components of PO(3) are given in Appendix B.

6.2 3×3 symmetric persymmetric

Given a nonzero symmetric persymmetric matrix $A = \left[\begin{smallmatrix} \alpha & \beta & \gamma \\ \beta & \delta & \beta \\ \gamma & \beta & \alpha \end{smallmatrix}\right]$, we want $W \in \mathrm{PO}_I(3)$ so that the $(1,2)$ element of WAW^T is zeroed out. Because such a similarity preserves symmetry as well as persymmetry, we will then have

$$WAW^T = \left[\begin{smallmatrix} * & 0 & \bullet \\ 0 & \times & 0 \\ \bullet & 0 & * \end{smallmatrix}\right]. \tag{6.2}$$

Using the parametrization $W = W(\theta)$ given in (6.1) and setting the $(1,2)$ element of WAW^T to zero yields

$$\tfrac{1}{\sqrt{2}}(\delta - \alpha - \gamma)cs + \beta(c^2 - s^2) = 0.$$

This equation is analogous to the one that arises in the solution of the 2×2 symmetric eigenproblem for the standard Jacobi method (see e.g., [26, p.350]), and it can be solved for (c, s) in exactly the same way. Let

$$\widehat{t} \;=\; \frac{2\sqrt{2}\,\beta}{\alpha + \gamma - \delta} \qquad \text{and} \qquad t \;=\; \frac{\widehat{t}}{1 + \sqrt{1 + \widehat{t}^2}}\,.$$

Then taking

$$(c, s) = \left(\frac{1}{\sqrt{1 + t^2}}\,,\, ct \right) \tag{6.3}$$

in (6.1) gives a $W = W(\theta)$ that achieves the desired form (6.2).

Algorithm 4 (3×3 **symmetric persymmetric).** *Given a symmetric persymmetric matrix* $A \in \mathbb{R}^{3\times 3}$*, this algorithm computes* $W \in \mathrm{PO}_I(3)$ *such that* WAW^T *is in canonical form as in* (6.2).

$$\begin{aligned}
&\widehat{t} = \frac{2\sqrt{2}\, a_{12}}{a_{11} + a_{13} - a_{22}}\,; \qquad t = \frac{\widehat{t}}{1 + \sqrt{1 + \widehat{t}^2}} \\
&c = \frac{1}{\sqrt{1+t^2}}\,; \qquad w_2 = \frac{1}{\sqrt{2}}\, ct \qquad\qquad \%\, s = \sqrt{2} w_2 \\
&w_1 = \tfrac{1}{2}(c+1)\,; \quad w_3 = \tfrac{1}{2}(c-1) \\
&W = \begin{bmatrix} w_1 & w_2 & w_3 \\ -w_2 & c & -w_2 \\ w_3 & w_2 & w_1 \end{bmatrix}
\end{aligned}$$

6.3 3×3 skew-symmetric persymmetric

Given a nonzero skew-symmetric persymmetric matrix $A = \left[\begin{smallmatrix} 0 & \beta & \alpha \\ -\beta & 0 & \beta \\ -\alpha & -\beta & 0 \end{smallmatrix}\right]$, we want $W \in \mathrm{PO}_I(3)$ so that the $(1,2)$ element of WAW^T is zeroed out. Because of the preservation of double structure, we will then have

$$WAW^T = \left[\begin{smallmatrix} 0 & 0 & \gamma \\ 0 & 0 & 0 \\ -\gamma & 0 & 0 \end{smallmatrix}\right]. \tag{6.4}$$

Proceeding as in section 6.2 leads to $\beta c - \frac{1}{\sqrt{2}} \alpha s = 0$. Among the two options for (c, s) satisfying this condition, the choice

$$(c, s) = \frac{1}{\sqrt{\alpha^2 + 2\beta^2}} \Big(|\alpha|\,, (\operatorname{sign} \alpha) \sqrt{2}\, \beta \Big) \tag{6.5}$$

corresponds to using the small angle for θ in the expression $W = W(\theta)$ given in (6.1), thus making W as close to the identity as possible.

Algorithm 5 (3×3 **skew-symmetric persymmetric).** *Given a skew-symmetric persymmetric matrix* $A \in \mathbb{R}^{3\times 3}$*, this algorithm computes* $W \in \mathrm{PO}_I(3)$ *such that* WAW^T *is in canonical form as in* (6.4).

$$\begin{aligned}
&\alpha = a_{13}\,; \qquad \beta = a_{12} \\
&\delta = \| \begin{bmatrix} \alpha & \beta & \beta \end{bmatrix} \|_2 \\
&c = \alpha/\delta\,; \qquad w_2 = \beta/\delta \qquad\qquad \%\, s = \sqrt{2}\, w_2 \\
&\text{if } \alpha < 0 \\
&\qquad c = -c\,; \qquad w_2 = -w_2 \\
&\text{endif} \\
&w_1 = \tfrac{1}{2}(c+1)\,; \qquad w_3 = \tfrac{1}{2}(c-1) \\
&W = \begin{bmatrix} w_1 & w_2 & w_3 \\ -w_2 & c & -w_2 \\ w_3 & w_2 & w_1 \end{bmatrix}
\end{aligned}$$

6.4 3×3 symmetric perskew-symmetric

Given a nonzero symmetric perskew-symmetric matrix $A = \left[\begin{smallmatrix} \alpha & \beta & 0 \\ \beta & 0 & -\beta \\ 0 & -\beta & -\alpha \end{smallmatrix}\right]$, we want $W \in \mathrm{PO}_I(3)$ so that

$$WAW^T = \left[\begin{smallmatrix} \gamma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\gamma \end{smallmatrix}\right]. \tag{6.6}$$

Since both perskew-symmetry and symmetry are automatically preserved by any similarity with $W \in \mathrm{PO}_I(3)$, we only need to ensure that the $(1,2)$ element of WAW^T is zero. This leads to the same condition as in section 6.3, that is, we need to choose the parameters c, s in $W = W(\theta)$ so that $\beta c - \frac{1}{\sqrt{2}}\alpha s = 0$. Consequently c, s chosen as in (6.5) yields $W \in \mathrm{PO}_I(3)$ as close to the identity as possible.

Algorithm 6 (3×3 symmetric perskew-symmetric). *Given a symmetric perskew-symmetric matrix $A \in \mathbb{R}^{3\times 3}$, this algorithm computes $W \in \mathrm{PO}_I(3)$ such that WAW^T is canonical form as in* (6.6).

$\alpha = a_{11}\ ; \qquad \beta = a_{12}$
$\delta = \| \begin{bmatrix} \alpha & \beta & \beta \end{bmatrix} \|_2$
$c = \alpha/\delta\ ; \qquad w_2 = \beta/\delta \qquad\qquad\qquad \%\, s = \sqrt{2}\, w_2$
if $\alpha < 0$
 $c = -c\ ; \qquad w_2 = -w_2$
endif
$w_1 = \frac{1}{2}(c+1)\ ; \qquad w_3 = \frac{1}{2}(c-1)$

$$W = \begin{bmatrix} w_1 & w_2 & w_3 \\ -w_2 & c & -w_2 \\ w_3 & w_2 & w_1 \end{bmatrix}$$

7 Perplectic orthogonal canonical forms

To build Jacobi algorithms from the 4×4 and 3×3 solutions described in sections 5 and 6, we need well-defined targets, that is, $n \times n$ structured canonical forms at which to aim our algorithms. The following theorem describes the canonical forms achievable by perplectic orthogonal (i.e. structure-preserving) similarities for each of the four classes of doubly-structured matrices under consideration.

Theorem 7.1. *Let $A \in \mathbb{R}^{n\times n}$.*

(a) *If A is symmetric and persymmetric then there exists a perplectic-orthogonal matrix P such that $P^{-1}AP$ is in structured "X-form", that is*

$$P^{-1}AP = \begin{bmatrix} a_1 & & 0 & & b_1 \\ & a_2 & & b_2 & \\ 0 & & \times & & 0 \\ & b_2 & & a_2 & \\ b_1 & & 0 & & a_1 \end{bmatrix}, \tag{7.1}$$

which is both symmetric and persymmetric.

(b) *If A is skew-symmetric and persymmetric, then there exists a perplectic-orthogonal matrix P such that $P^{-1}AP$ is antidiagonal and skew-symmetric, that is,*

$$P^{-1}AP = \begin{bmatrix} 0 & & & & -a \\ & & & -b & \\ & & \iddots & & \\ & b & & & \\ a & & & & 0 \end{bmatrix}. \tag{7.2}$$

(c) *If A is symmetric and perskew-symmetric, then there exists a perplectic-orthogonal matrix P such that $P^{-1}AP$ is diagonal and perskew-symmetric, that is,*

$$P^{-1}AP = \begin{bmatrix} a & & & & 0 \\ & b & & & \\ & & \ddots & & \\ & & & -b & \\ 0 & & & & -a \end{bmatrix}. \tag{7.3}$$

(d) *If A is skew-symmetric and perskew-symmetric then there exists a perplectic-orthogonal matrix P such that $P^{-1}AP$ has the following "block X-form",*

$$P^{-1}AP = \begin{bmatrix} A_1 & & & & & & B_1 \\ & A_2 & & 0 & & B_2 & \\ & & \ddots & & \iddots & & \\ 0 & & & Z & & & 0 \\ & & \iddots & & \ddots & & \\ & -B_2 & & 0 & & -A_2 & \\ -B_1 & & & & & & -A_1 \end{bmatrix}, \tag{7.4}$$

where A_i and B_i are 2×2 real matrices of the form $\left[\begin{smallmatrix} 0 & a_i \\ -a_i & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} b_i & 0 \\ 0 & -b_i \end{smallmatrix}\right]$, respectively, and

$$Z = \begin{cases} \emptyset & \textit{if } n \equiv 0 \mod 4, \\ 0 & \textit{if } n \equiv 1 \mod 4, \\ \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right] & \textit{if } n \equiv 2 \mod 4, \\ \left[\begin{smallmatrix} 0 & -c & 0 \\ c & 0 & c \\ 0 & -c & 0 \end{smallmatrix}\right] & \textit{if } n \equiv 3 \mod 4. \end{cases}$$

Since similarity by P preserves structure, the block X-form given by (7.4) *is both skew-symmetric and perskew-symmetric.*

The result of part(a) cannot be improved, as the matrix $A = I + R$ demonstrates: it is symmetric and persymmetric, and impervious to any perplectic orthogonal similarity. The result of part(d) is also the best that can be achieved: by the discussion accompanying (5.1), the 4×4 skew-symmetric perskew-symmetric matrices $\left[\begin{smallmatrix} A_i & B_i \\ -B_i & -A_i \end{smallmatrix}\right]$ cannot be reduced further. The main impetus for conjecturing the canonical forms given in Theorem 7.1 comes from the quaternion solution in the case when $n = 4$. For a proof of the general case, see [16]. Complex canonical forms for various classes of doubly structured matrices in $\mathbb{C}^{n\times n}$ have been discussed in [1], [20]. However, the real canonical forms given by Theorem 7.1 cannot be readily derived from the results in [1], [20].

The next section presents structure-preserving Jacobi algorithms to achieve the canonical forms in (7.1)-(7.3). As was remarked earlier, a consequence of (5.1) is that a Jacobi algorithm for doubly skewed matrices cannot be built using 4×4 subproblems as a basis. Finding a structure-preserving algorithm to achieve the canonical form given in (7.4) remains an open problem.

8 Sweep design

For a Jacobi algorithm to have a good rate of convergence to the desired canonical form, it is essential that every element of the $n\times n$ matrix be part of a target subproblem at least once during a sweep, whether the sweep is cyclic or quasi-cyclic. There are several ways to design a sequence of *structured* subproblems that give rise to such sweeps.

Let $A = \begin{bmatrix} B & x & C \\ y^T & \alpha & z^T \\ D & w & E \end{bmatrix} \in \mathbb{R}^{n\times n}$ have symmetry or skew-symmetry about the main diagonal as well as the anti-diagonal. Here B, C, D, $E \in \mathbb{R}^{m\times m}$, where $m = \lfloor \frac{n}{2} \rfloor$. If n is odd, then x, y, z, $w \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$; otherwise these variables are absent.

First note that an off-diagonal element a_{ij} chosen from the $m\times m$ block B uniquely determines a 4×4 principal submatrix $A_4[i,j]$ that is *centrosymmetrically embedded* in A; this means that $A_4[i,j]$ is located in rows and columns i, j, $n-j+1$ and $n-i+1$:

$$A_4[i,j] = \left[\begin{array}{cc|cc} a_{ii} & a_{ij} & a_{i,n-j+1} & a_{i,n-i+1} \\ a_{ji} & a_{jj} & a_{j,n-j+1} & a_{j,n-i+1} \\ \hline a_{n-j+1,i} & a_{n-j+1,j} & a_{n-j+1,n-j+1} & a_{n-j+1,n-i+1} \\ a_{n-i+1,i} & a_{n-i+1,j} & a_{n-i+1,n-j+1} & a_{n-i+1,n-i+1} \end{array}\right]. \tag{8.1}$$

Centrosymmetrically embedded submatrices inherit both structures from the parent matrix A — symmetry or skew-symmetry together with persymmetry or perskew-symmetry. Furthermore, when n is even, *any* cyclic or quasi-cyclic sweep of the block B consisting of 2×2 principal submatrices will generate a corresponding cyclic (respectively quasi-cyclic) sweep of A, comprised entirely of 4×4 centrosymmetrically embedded subproblems. An illustration when $n = 8$ is given in Figure 8.1 using a row-cyclic sweep for B. The entry denoted by $\diamond$ determines the position of the rest of the elements in the current target subproblem. These are represented by heavy bullets. Observe that every entry of A is part of a target submatrix during the course of the sweep, and that this property will hold for any choice of a 2×2 based cyclic or quasi-cyclic sweep pattern for B. Animated views, in various formats, of a row-cyclic sweep on a 12×12 matrix can be found at `http://www.cscamm.umd.edu/~ddunlavy/perplectic.html`.

When n is odd, a sweep will involve centrosymmetrically embedded 3×3 targets as well as 4×4 ones. A 3×3 target $A_3[i]$ is determined by a single element a_{ii} chosen from

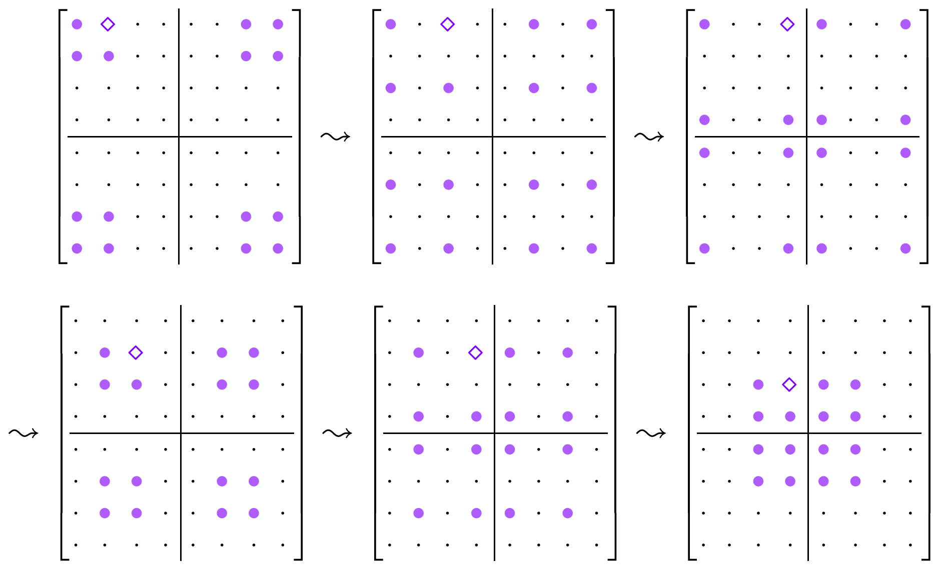

Figure 8.1: Row-cyclic structured sweep, $n = 8$

the $m \times m$ block B, and always involves elements from x, w, y^T, z^T, and the center element $\alpha = a_{m+1,m+1}$:

$$A_3[i] = \begin{bmatrix} a_{ii} & a_{i,m+1} & a_{i,n-i+1} \\ a_{m+1,i} & a_{m+1,m+1} & a_{m+1,n-i+1} \\ a_{n-i+1,i} & a_{n-i+1,m+1} & a_{n-i+1,n-i+1} \end{bmatrix}. \tag{8.2}$$

Animated views, in various formats, of a row-cyclic sweep on a 13×13 matrix can be found at `http://www.cscamm.umd.edu/~ddunlavy/perplectic.html`. Figure 8.2 illustrates such a sweep for $n = 7$; entries in locations corresponding to x, y, z^T, w^T and α are depicted by $*$.

Once a target submatrix of A has been identified, $W \in \mathrm{PO}_I(4)$ or $W \in \mathrm{PO}_I(3)$ is constructed using the appropriate algorithm from section 5.1, 5.2 or 5.3, or section 6.2, 6.3 or 6.4. Centrosymmetrically embedding W into I_n yields a matrix in $\mathrm{PO}_I(n)$.

A Jacobi algorithm built on these ideas is illustrated in Algorithm 7 for a symmetric persymmetric matrix A, using a row-cyclic ordering. Since in this case A is being driven to X-form as in (7.1),

$$\mathrm{off}(A) = \sqrt{\sum_{(i,j)\in\mathcal{S}} a_{ij}^2} \quad \text{where } \mathcal{S} = \{(i,j) : 1 \le i,j \le n, j \ne i, j \ne n-i+1\}$$

is used as a measure of the deviation from the desired canonical form.

Figure 8.3 depicts a slide show of Algorithm 7 running on a 12×12 symmetric persymmetric matrix. A snapshot of the matrix is taken after each iteration, that is, after each 4×4 similarity transformation. Each row of snapshots shows the progression during a sweep. In this case, the algorithm terminates after 5 sweeps. Movies of Algorithm 7

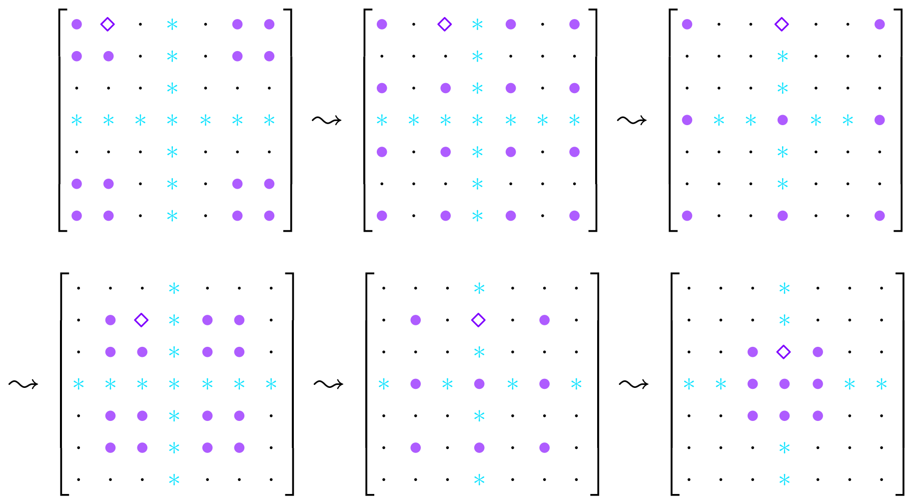

Figure 8.2: Row-cyclic structured sweep, $n = 7$

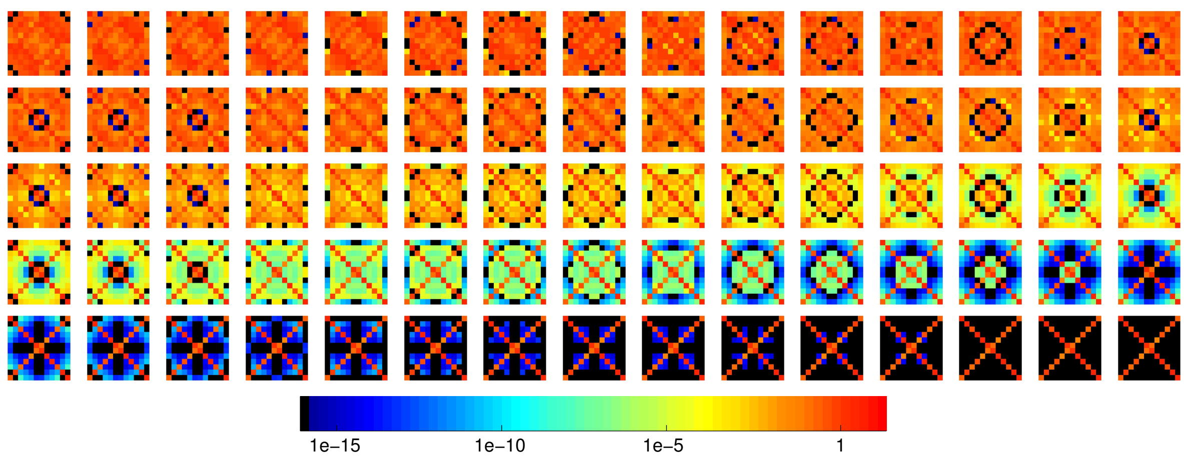

Figure 8.3: Algorithm 7 running on a 12×12 symmetric persymmetric matrix

running on 12×12 and 32×32 symmetric persymmetric matrices can be downloaded from `http://www.cscamm.umd.edu/~ddunlavy/perplectic.html`.

Algorithm 7 (Row-cyclic Jacobi for symmetric persymmetric matrices). *Given a symmetric persymmetric matrix* $A \in \mathbb{R}^{n\times n}$, *and a tolerance* $\mathrm{tol} > 0$, *this algorithm overwrites* A *with its approximate canonical form* PAP^T *where* $P \in \mathrm{PO}_I(n)$ *and* $\mathrm{off}(PAP^T) < \mathrm{tol}\|A\|_F$. *The matrix* P *is also computed.*

$P = I_n\,;\quad \delta = \mathrm{tol}\,\|A\|_F\,;\quad m = \lfloor n/2 \rfloor$
while $\mathrm{off}(A) > \delta$
  for $i = 1\colon m-1$
    for $j = i+1\colon m$
      Use Algorithm 1 to find $W \in \mathbb{R}^{4\times 4}$ such that $A_4[i,j]$ is in X-form

$\widehat{P} = I_n\,; \quad \widehat{P}_4[i,j] = W$
$A = \widehat{P} A \widehat{P}^T$
$P = \widehat{P} P$
endfor
if n is odd then
Use Algorithm 4 to find $W \in \mathbb{R}^{3\times 3}$ such that $A_3[i]$ is in X-form
$\widehat{P} = I_n\,; \quad \widehat{P}_3[i] = W$
$A = \widehat{P} A \widehat{P}^T$
$P = \widehat{P} P$
endif
endfor
if n is odd then
Use Algorithm 4 to find $W \in \mathbb{R}^{3\times 3}$ such that $A_3[m]$ is in X-form
$\widehat{P} = I_n\,; \quad \widehat{P}_3[m] = W$
$A = \widehat{P} A \widehat{P}^T$
$P = \widehat{P} P$
endif
endwhile % A is now in canonical form as in (7.1)

Parallelizable Jacobi orderings in the 2×2 setting (see for example [7], [8], [14], [18], [19], [24]) on the $m \times m$ block B yield corresponding parallelizable structure-preserving sweeps for the $n \times n$ matrix A. Finally we note that since the double structure of the $n \times n$ matrix is always preserved, both storage requirements and operation counts can be lowered by roughly a factor of four.

9 Numerical Results

We present a brief set of numerical experiments to demonstrate the effectiveness of our algorithms. All computations were done using MATLAB Version 5.3.0 on a Sun Ultra 5 with IEEE double-precision arithmetic and machine precision $\epsilon = 2.2204 \times 10^{-16}$. As stopping criteria we chose $\mathrm{reloff}(A) < \mathrm{tol}$, where $\mathrm{reloff}(A) = \mathrm{off}(A)/\|A\|_F$. Here $\mathrm{off}(A)$ is the appropriate off-diagonal norm for the structure under consideration, $\|A\|_F$ is the Frobenius norm of A, and $\mathrm{tol} = \epsilon\|A\|_F$.

For each of the three doubly-structured classes, and for each $n = 20, 25, \ldots, 100$, the algorithms were run on 100 random $2n \times 2n$ structured matrices with entries normally distributed with mean zero ($\mu = 0$) and variance one ($\sigma = 1$). The tests were repeated for matrices with entries uniformly distributed on the interval $[-1, 1]$ with no significant differences in the results. The results are reported in Figures 9.1-9.2 and Tables 9.1-9.3 and discussed below.

- The methods always converged, and the off-norm always decreased monotonically. The convergence rate was initially linear, but asymptotically quadratic. This is shown in Figure 9.1 using a sample 200×200 matrix from each of the three classes.

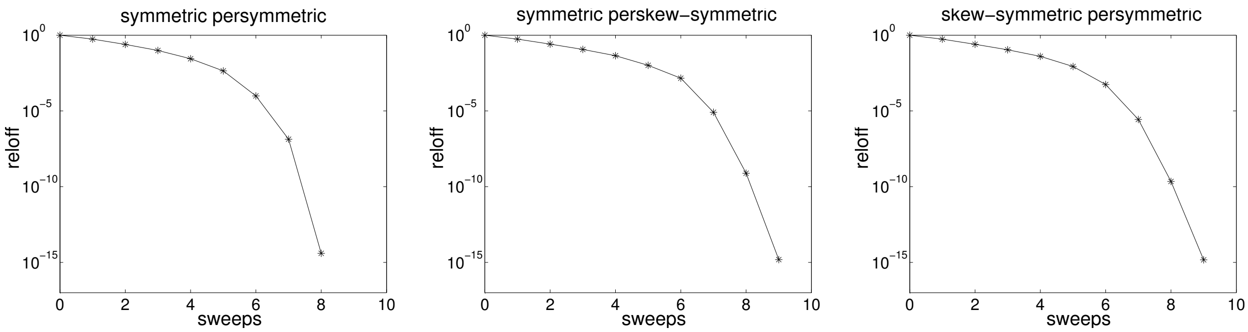

Figure 9.1: Typical convergence behavior of 200×200 matrices

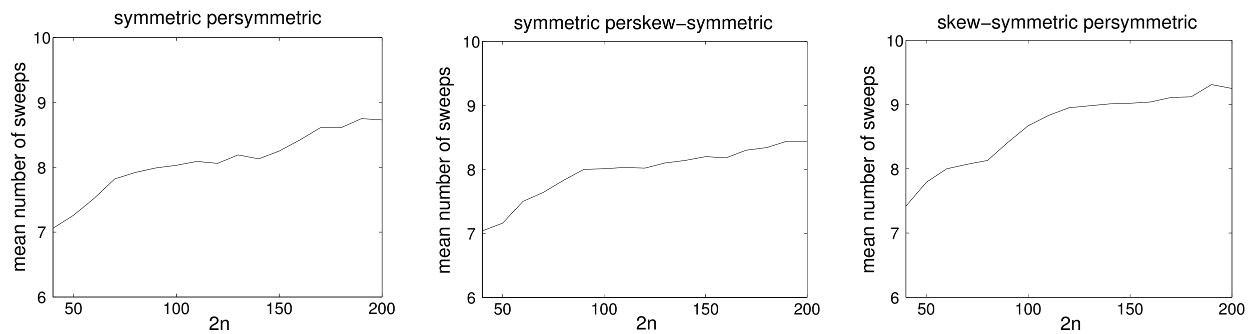

Figure 9.2: Average number of sweeps for convergence for $2n \times 2n$ matrices

- It was experimentally observed that the number of sweeps needed for convergence depends only on matrix size: the standard deviation of the average number of sweeps was consistently very low — between 0 and 0.52. Figure 9.2 suggests that roughly $O(\log n)$ sweeps suffice. This leads to an *a priori* stopping criterion, which is an important consideration on parallel architectures: a stopping criteria that depends on global knowledge of the matrix elements would undermine the advantage gained by parallelism.

- As the matrices are always either symmetric or skew-symmetric, all eigenvalues have condition number equal to 1, are all real or pure imaginary, and can be easily sorted and compared with the eigenvalues computed by MATLAB's eig function. The maximum relative error, releig $= \max_j |\lambda_j^{eig} - \lambda_j^{jac}|/|\lambda_i^{eig}|$ was of the order 10^{-13} as shown in the last column of Tables 9.1-9.3.

- The computed perplectic orthogonal transformations P from which the eigenvectors or invariant subspaces can be obtained were both perplectic as well as orthogonal to within 6.3×10^{-14}, as measured by $\|P^T RP - R\|$ and $\|P^T P - I\|$ in Tables 9.1-9.3. Since perplectic orthogonal matrices are centrosymmetric (see section 3), the deviation from centrosymmetric block structure $[\begin{smallmatrix} U & V \\ RVR & RUR \end{smallmatrix}]$ can be measured by block $= \|P(1:n, 1:n) - RP(n+1:2n, n+1:2n)R\|_F + \|P(1:$

$2n$	sweeps	reloff	$\|P^TRP - R\|_F$	$\|P^TP - I\|_F$	block	releig
50	7.22	4.04×10^{-16}	1.40×10^{-14}	1.42×10^{-14}	3.03×10^{-15}	3.29×10^{-14}
100	8.02	4.66×10^{-16}	2.98×10^{-14}	3.00×10^{-14}	4.55×10^{-15}	1.02×10^{-13}
150	8.27	4.09×10^{-15}	4.50×10^{-14}	4.52×10^{-14}	5.76×10^{-15}	1.47×10^{-13}
200	8.84	1.99×10^{-15}	6.22×10^{-14}	6.25×10^{-14}	6.77×10^{-15}	1.09×10^{-13}

Table 9.1: $2n \times 2n$ symmetric persymmetric matrices

$2n$	sweeps	reloff	$\|P^TRP - R\|_F$	$\|P^TP - I\|_F$	block	releig
50	7.10	1.02×10^{-15}	9.79×10^{-15}	9.95×10^{-15}	3.01×10^{-15}	3.30×10^{-14}
100	8.02	1.27×10^{-15}	1.99×10^{-14}	2.01×10^{-14}	4.55×10^{-15}	6.06×10^{-14}
150	8.14	3.16×10^{-15}	2.75×10^{-14}	2.78×10^{-14}	5.68×10^{-15}	8.60×10^{-14}
200	8.54	6.18×10^{-15}	3.82×10^{-14}	3.84×10^{-14}	6.69×10^{-15}	1.30×10^{-13}

Table 9.2: $2n \times 2n$ symmetric perskew-symmetric matrices

$2n$	sweeps	reloff	$\|P^TRP - R\|_F$	$\|P^TP - I\|_F$	block	releig
50	7.84	1.03×10^{-15}	1.08×10^{-14}	1.10×10^{-14}	3.18×10^{-15}	1.68×10^{-14}
100	8.67	2.25×10^{-15}	2.25×10^{-14}	2.27×10^{-14}	4.77×10^{-15}	8.22×10^{-14}
150	9.05	2.52×10^{-15}	3.21×10^{-14}	3.24×10^{-14}	6.00×10^{-15}	7.05×10^{-14}
200	9.28	4.44×10^{-15}	4.26×10^{-14}	4.28×10^{-14}	7.03×10^{-15}	1.11×10^{-13}

Table 9.3: $2n \times 2n$ skew-symmetric persymmetric matrices

$n, n+1 : 2n) - RP(n+1 : 2n, 1 : n)R\|_F$; both terms in this sum had about the same size. Note that eigenvectors computed by MATLAB's eig function cannot be directly compared to the perplectic bases obtained by our algorithms.

10 Concluding remarks

We have presented new structured canonical forms for matrices that are symmetric or skew-symmetric with respect to the main diagonal as well as the anti-diagonal, and developed structure-preserving Jacobi algorithms to compute these forms in three out of four cases. In the fourth case – when the matrix is skew-symmetric with respect to both diagonals – a structure preserving method to compute the corresponding canonical form remains an open problem.

In order to effectively design structure preserving transformations for our algorithms, explicit parametrizations of the perplectic orthogonal groups PO(3) and PO(4) were developed. These groups are disconnected, so in order to promote good convergence behavior, the algorithms were designed to accomplish their goals using only transformations in the connected component of the identity matrix.

In addition to preserving the double structure in the parent matrix throughout the computation, these algorithms are inherently parallelizable and are experimentally observed to be asymptotically quadratically convergent. It is expected that the recent analysis by Tisseur [25] of a related family of algorithms can also be applied to this work to show that these methods are not only backward stable, but in fact strongly backward stable.

A The Quaternion Basis for $\mathbb{R}^{4\times 4}$

$$\underset{1\otimes 1}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}} \quad \underset{1\otimes i}{\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}} \quad \underset{1\otimes j}{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}} \quad \underset{1\otimes k}{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}}$$

$$\underset{i\otimes 1}{\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}} \quad \underset{i\otimes i}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}} \quad \underset{i\otimes j}{\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}} \quad \underset{i\otimes k}{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}$$

$$\underset{j\otimes 1}{\begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}} \quad \underset{j\otimes i}{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}} \quad \underset{j\otimes j}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}} \quad \underset{j\otimes k}{\begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}$$

$$\underset{k\otimes 1}{\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}} \quad \underset{k\otimes i}{\begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}} \quad \underset{k\otimes j}{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}} \quad \underset{k\otimes k}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}$$

B Parametrizations of $\mathrm{PO}(3)$ and $\mathrm{PO}(4)$

Since the only 2×2 matrices that are centrosymmetric and orthogonal are $\pm I_2$ and $\pm R_2$, $\mathrm{PO}(2)$ is a discrete group with four connected components. The explicit parametrizations of $\mathrm{PO}(3)$ and $\mathrm{PO}(4)$ developed here show that each of these groups also has exactly four connected components.

B.1 $\mathrm{PO}(3)$

Let $W \in \mathrm{PO}(3)$. By (3.1), W is centrosymmetric and hence can be expressed as

$$W = \begin{bmatrix} \alpha & \beta & \gamma \\ \delta & \epsilon & \delta \\ \gamma & \beta & \alpha \end{bmatrix}.$$

Using orthogonality we get

$$\alpha^2 + \delta^2 + \gamma^2 = \alpha^2 + \beta^2 + \gamma^2 \implies \delta = \pm\beta$$

If $\delta = 0$, then $\beta = 0$ and $\epsilon = \pm 1$. This means $\left[\begin{smallmatrix}\alpha & \gamma \\ \gamma & \alpha\end{smallmatrix}\right] \in \mathrm{PO}(2)$, and hence $\left[\begin{smallmatrix}\alpha & \gamma \\ \gamma & \alpha\end{smallmatrix}\right]$ is $\pm I_2$ or $\pm R_2$. Otherwise,

$$\alpha\beta + \delta\epsilon + \gamma\beta = 0 \implies \delta\epsilon = -\beta(\alpha+\gamma) \implies \epsilon = \begin{cases} -(\alpha+\gamma) & \text{if } \delta = \beta, \\ \alpha+\gamma & \text{if } \delta = -\beta \end{cases}$$

$$2\alpha\gamma + \beta^2 = 0, \;\; \alpha^2 + \beta^2 + \gamma^2 = 1 \implies (\alpha+\gamma)^2 + 2\beta^2 = 1.$$

Thus we may write $\alpha + \gamma = \cos\theta$ and $\beta = \frac{1}{\sqrt{2}}\sin\theta$, where $\theta \in [0, 2\pi)$. Substituting for β in $2\alpha\gamma + \beta^2 = 0$ yields $4\alpha\gamma = -\sin^2\theta$. Consequently,

$$\alpha + \gamma = \cos\theta \implies 4\alpha^2 - \sin^2\theta = 4\alpha\cos\theta \implies \alpha = \tfrac{1}{2}(\cos\theta \pm 1), \;\; \gamma = \tfrac{1}{2}(\cos\theta \mp 1)$$

This gives us a parametrization of PO(3) that reveals this group has four connected components, two of which consist of perplectic orthogonals with positive determinant (that is, perplectic orthogonal rotations). Using the abbreviations $c = \cos\theta$, $s = \sin\theta$, the connected component containing the identity is given by

$$\mathrm{PO}_I(3) = \left\{ W(\theta) = \frac{1}{2}\begin{bmatrix} c+1 & \sqrt{2}s & c-1 \\ -\sqrt{2}s & 2c & -\sqrt{2}s \\ c-1 & \sqrt{2}s & c+1 \end{bmatrix}, \text{where } \theta \in [0, 2\pi) \right\}. \tag{B.1}$$

Each $W(\theta)$ represents a rotation by angle θ about the axis through $\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$. As is the case for the connected component of the identity in any Lie group, $\mathrm{PO}_I(3)$ is a normal subgroup of PO(3). The other component containing perplectic rotations is parametrized by

$$\frac{1}{2}\begin{bmatrix} c-1 & \sqrt{2}s & c+1 \\ \sqrt{2}s & -2c & \sqrt{2}s \\ c+1 & \sqrt{2}s & c-1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} W(\theta),$$

while the two components containing perplectic orthogonals with negative determinant are given by

$$\frac{1}{2}\begin{bmatrix} c+1 & \sqrt{2}s & c-1 \\ \sqrt{2}s & -2c & \sqrt{2}s \\ c-1 & \sqrt{2}s & c+1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} W(\theta),$$

and

$$\frac{1}{2}\begin{bmatrix} c-1 & \sqrt{2}s & c+1 \\ -\sqrt{2}s & 2c & -\sqrt{2}s \\ c+1 & \sqrt{2}s & c-1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} W(\theta).$$

Thus these three connected components correspond to left cosets of the normal subgroup $\mathrm{PO}_I(3)$. Observe that these parametrizations also show that each of the components of PO(3) is homeomorphic to the circle S^1.

B.2 PO(4)

From section 4.3 we have the quaternion parametrizations

$$\mathrm{PO}_I(4) = \Big\{ u \otimes v : |u| = |v| = 1, u \in \mathrm{span}\{1, j\}, v \in \mathrm{span}\{1, i\} \Big\}$$

for the connected component containing the identity, and

$$\Big\{ u \otimes v : |u| = |v| = 1, u \in \mathrm{span}\{i, k\}, v \in \mathrm{span}\{j, k\} \Big\} \tag{B.2}$$

for the other connected component of PO(4) containing rotations. Writing $u = \cos\alpha + (\sin\alpha)j$ and $v = \cos\beta + (\sin\beta)i$ where $0 \le \alpha < 2\pi$ and $0 \le \beta < 2\pi$, we can write $W(\alpha, \beta) \in \mathrm{PO}_I(4)$ in matrix form as

$$\begin{aligned}
W(\alpha,\beta) = \;& u \otimes v \;=\; (u \otimes 1)(1 \otimes v) \\
= & \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha & 0 \\ 0 & \cos\alpha & 0 & \sin\alpha \\ \sin\alpha & 0 & \cos\alpha & 0 \\ 0 & -\sin\alpha & 0 & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta & 0 & 0 \\ -\sin\beta & \cos\beta & 0 & 0 \\ 0 & 0 & \cos\beta & -\sin\beta \\ 0 & 0 & \sin\beta & \cos\beta \end{bmatrix} \\
= & \begin{bmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta & -\sin\alpha\cos\beta & \sin\alpha\sin\beta \\ -\cos\alpha\sin\beta & \cos\alpha\cos\beta & \sin\alpha\sin\beta & \sin\alpha\cos\beta \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta & \cos\alpha\cos\beta & -\cos\alpha\sin\beta \\ \sin\alpha\sin\beta & -\sin\alpha\cos\beta & \cos\alpha\sin\beta & \cos\alpha\cos\beta \end{bmatrix}.
\end{aligned}$$

A perplectic rotation in the connected component given by (B.2) can then be expressed as

$$(k \otimes k) \cdot W(\alpha,\beta) \;=\; \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{bmatrix} W(\alpha,\beta)\,.$$

There are two more connected components of PO(4), containing matrices with negative determinant. They are given by the parametrizations

$$\left\{ \begin{bmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{bmatrix} W(\alpha,\beta) \right\}, \qquad \left\{ \begin{bmatrix} 1 & & & \\ & 0 & -1 & \\ & -1 & 0 & \\ & & & 1 \end{bmatrix} W(\alpha,\beta) \right\}.$$

Once again, the connected component containing the identity is a normal subgroup of PO(4); the parametrizations for the other three connected components show that they are cosets of $\mathrm{PO}_I(4)$.

References

[1] Gregory Ammar, Christian Mehl, and Volker Mehrmann. Schur-like forms for matrix Lie groups, Lie algebras and Jordan algebras. *Linear Algebra Appl.*, 287:11–39, 1999.

[2] Theodor Bröcker and Tammo tom Dieck. *Representations of Compact Lie Groups*. Springer-Verlag, New York, 1985.

[3] R. Byers. A Hamiltonian-Jacobi algorithm. *IEEE Trans. Automat. Contr.*, AC-35:566–570, 1990.

[4] Arthur Cayley. On certain results relating to quaternions. In *The Collected Mathematical Papers*, volume 1, pages 123–126. Johnson Reprint Company, New York, 1963. Originally published in The Philos. Mag., **26** (1845), pp. 141–145.

[5] Arthur Cayley. Recherches ultérieures sur les déterminants gauches. In *The Collected Mathematical Papers*, volume 2, pages 202–215. Johnson Reprint Company, New York, 1963. Originally published in J. Reine Angew. Math., **50** (1855), pp. 299-313.

[6] H. D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, and R. Remmert. *Numbers*. Springer-Verlag, New York, 1991.

[7] P. J. Eberlein. On using the Jacobi method on the hypercube. In M. T. Heath, editor, *Proc. 2nd Conference on Hypercube Multiprocessors*, pages 605–611. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1987.

[8] P. J. Eberlein and Haesun Park. Efficient implementation of Jacobi algorithms and Jacobi sets on distributed memory architectures. *J. Parallel Dist. Comput.*, 8:358–366, 1990.

[9] H. Faßbender, D. S. Mackey, N. Mackey, and H. Xu. Real and complex Hamiltonian square roots of skew-Hamiltonian matrices. Technical Report 92, Department of Mathematics & Statistics, Western Michigan University, 1999. Updated version of [11].

[10] Heike Faßbender, D. Steven Mackey, and Niloufer Mackey. Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian problems. *Linear Algebra Appl.*, 332-334:37–80, 2001.

[11] Heike Faßbender, D. Steven Mackey, Niloufer Mackey, and Hongguo Xu. Hamiltonian square roots of skew-Hamiltonian matrices. *Linear Algebra Appl.*, 287:125–159, 1999.

[12] Derek Hacon. Jacobi's method for skew-symmetric matrices. *SIAM J. Matrix Anal. Appl.*, 14(3):619–628, 1993.

[13] William R. Hamilton. On quaternions, or a new system of imaginaries in algebra; with some geometrical illustrations. In H. Halberstam and R. E. Ingram, editors, *The Collected Mathematical Papers*, volume 3, pages 355–362. Cambridge University Press, Cambridge, 1967. Originally published in *Proc. Royal Irish Academy* **3** (1847), pp. 1–16; communicated November 1844.

[14] F. T. Luk and H. Park. On parallel Jacobi orderings. *SIAM J. Sci. Stat. Comp.*, 10:18–26, 1989.

[15] D. Steven Mackey and Niloufer Mackey. Canonical forms for doubly structured matrices, part 1: Real symplectic case. Technical report, Department of Mathematics & Statistics, Western Michigan University. In Preparation.

[16] D. Steven Mackey and Niloufer Mackey. Canonical forms for doubly structured matrices, part 2: Real perplectic case. Technical report, Department of Mathematics & Statistics, Western Michigan University. In Preparation.

[17] Niloufer Mackey. Hamilton and Jacobi meet again: Quaternions and the eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 16:421–435, April 1995.

[18] Mythili Mantharam and P. J. Eberlein. Block recursive algorithm to generate Jacobi-sets. *Parallel Computing*, 19:481–496, 1993.

[19] Walter F. Mascarenhas. *On the convergence of the Jacobi method for arbitrary orderings.* PhD thesis, Massachusetts Institute of Technology, Cambridge, 1991.

[20] Christian Mehl, Volker Mehrmann, and Hongguo Xu. Canonical forms for doubly structured matrices and pencils. *Electronic Journal of Linear Algebra*, 7:112–151, 2000.

[21] Richard S. Pierce. *Associative Algebras.* Springer-Verlag, New York, 1982.

[22] Elmer G. Rees. *Notes on Geometry.* Springer–Verlag, Berlin, 1983.

[23] Russell M. Reid. Some eigenvalue properties of persymmetric matrices. *SIAM Rev.*, 39(2):313–316, 1997.

[24] A. H. Sameh. On Jacobi and Jacobi-like algorithms for a parallel computer. *Math. of Comp.*, 25:579–590, 1971.

[25] Françoise Tisseur. Stability of structured Hamiltonian eigensolvers. *SIAM J. Matrix Anal. Appl.*, 23(1):103–125, 2001.

[26] David S. Watkins. *Fundamentals of Matrix Computations.* John Wiley, New York, 1991.